

基于 BP 神经网络的传感器网络动态采样模型研究

阒 杰^{1,2,3}, 张瑞瑞^{1,2,4}, 陈立平^{1,2}, 徐 刚^{1,2}

(1. 北京农业智能装备技术研究中心, 北京 100097; 2. 国家农业智能装备工程技术研究中心, 北京 100097;
3. 首都师范大学 信息工程学院, 北京 100048; 4. 中国农业大学 信息与电气工程学院, 北京 100083)

摘要: 能耗控制对于农业环境监测无线传感器网络系统具有重要意义; 基于误差反向传播的多层前馈神经网络预测和阈值分析建立了一种土壤温度传感器网络动态采样模型, 实现了基于土壤温度周期变化特征的采样频率实时调整方法, 达到减少网络冗余数据, 降低网络功耗的目的; 以环境温度和空气相对湿度为 BP 神经网络实测输入, 土壤温度为预测输出, 通过判断输出是否进入阈值区间动态调整采样周期; 仿真实验结果表明, 对于具有周期性特点的土壤温度, BP 网络模型对其预测值和实测值之间的均方根误差 RMSE 及绝对误差 AE 分别为 0.83 °C 和 0.54 °C; 相比于连续采样, 阈值分析动态采样次数减少 30%。

关键词: 无线传感器网络; BP 神经网络; 动态调度; 能量优化

Dynamic Sampling Model Research of Sensor Network Based on BP Neural Network

Kan Jie^{1,2,3}, Zhang Ruirui^{1,2,4}, Chen Liping^{1,2}, Xu Gang^{1,2}

(1. Beijing Research Center of Intelligent Equipment for Agriculture, Beijing 100097, China;
2. National Research Center of Intelligent Equipment for Agriculture, Beijing 100097, China;
3. College of Information Engineering, Capital Normal University, Beijing 100048, China;
4. College of Information and Electrical Engineering, China Agriculture University, Beijing 100083, China)

Abstract: Energy consumption control for agricultural environment monitoring wireless sensor network system is of great significance. A dynamic sampling model of soil temperature sensor network based on BP neural network prediction and threshold analysis, for real-time adjustment of sampling frequency in the light of the periodic variation features of soil temperature, was constructed to reduce the redundant sampling data and power consumption of the network. Using environment temperature and air humidity as the actual measure input of the BP neural network, soil temperature as predict output of the BP neural network, we adjust the sampling periods dynamically by observing whether the output of the BP neural network is entering the threshold interval. The simulation results show that, for the periodic soil temperature, the root mean squared error (RMSE) and absolute error (AE) between predicted values and measured values were 0.83 °C and 0.54 °C. Comparing to continuous sampling, dynamic sampling reduce the data collection by 30%.

Keywords: wireless sensor network; BP neural network; dynamic scheduling; energy optimization

0 引言

无线传感器网络 (wireless sensor network, WSN) 是以感知、采集和处理网络覆盖区域中被感知对象的信息为目的, 利用大量传感器节点以无线、多跳、自组织方式构成的数据协同处理网络^[1]。如何最大限度地挖掘网络有限的资源、延长网络生命周期依然是传感器网络研究的关键问题之一^[2-5]。

基于数据预测的传感器网络动态采样方法是一种动态采样主要的实现方式。文献 [6] 提出了一种基于感知数据概率模型的无线传感器采样调度方法, 使用简单的 AR (P) 模型进

行数据预测, 通过与实际采样值之间的误差分析动态调整节点采样频率。文献 [7-8] 提出一种适用于无线传感器网络土壤湿度测量的新型动态采样控制策略, 将土壤湿度量化为马尔可夫过程, 利用部分可观察马尔可夫决策问题, 通过动态远程控制传感器节点的采样周期, 以较少的采样点满足测量需求, 降低数据冗余, 减少网内数据传输量。文献 [9] 提出了一种基于马尔科夫链的无线传感器网络分布式调度算法, 通过构建节点状态转换马尔科夫概率模型, 动态修正节点数据包发送概率和节点休眠概率, 达到节省能量的目的。以上节点调度方法主要都是在时间序列维度下对单一环境参数的分析建模, 基于多参数相关性的动态采样建模研究则相对较少, 且上述算法较为复杂, 对于计算能力受限的传感器节点, 有较大实施难度。

依据参数变化规律, 动态调整采样周期, 减少网络数据冗余是延长网络生命周期的重要方法。本文基于误差反向传播的多层前馈神经网络 (back propagation neural network) 预测和阈值分析建立了一种土壤温度传感器网络动态采样模型, 并对模型的应用效果进行了仿真实验分析。

收稿日期: 2014-10-08; 修回日期: 2014-12-16。

基金项目: 北京市自然科学基金资助项目 (4121001); 国家高技术研究发展计划资助项目 (2012AA101901); 北京市农林科学院青年科研基金资助项目 (QNJJ201217)。

作者简介: 阒 杰 (1989-), 男, 安徽滁州人, 主要从事无线传感器网络及数据分析方向的研究。

通讯作者: 陈立平 (1973-), 女, 博士, 研究员, 主要从事农业信息技术及智能装备技术方向的研究。

1 神经网络模型

1.1 BP 神经网络结构

BP 网络是一种单向传播的多层前向网络，由输出层向输入层逐步反馈的学习算法，误差向后传播。已证明，仅有一个非线性隐藏层的 BP 前馈网络就能以任意精度逼近任意复杂度的函数^[10-11]。在数值预测方面，BP 神经网络不需要预先确定样本数据的数学模型，仅通过学习样本数据即可进行精确预测。神经网络的优点在于其良好的非线性映射能力，可以方便灵活的对成因复杂的未知系统进行高精度建模，解决了传统预测模型难以处理高维非线性以及缺少自适应和自学习能力的问题^[12-13]。

如图 1 所示，三层 BP 网络模型包括输入层、隐藏层和输出层。输入层有 M 个节点 X_i ($i=1, 2, \dots, m$)，接收预处理后的参考矢量值 (X_1, \dots, X_m)；中间隐藏层有 L 个处理单元节点，激活函数可选择曲线 sigmoid 函数， $W_{ij}^{(1)}$ ($i=1, \dots, m; j=1, \dots, l$) 为输入层和隐藏层的连接权值；输出层有 N 个节点 Y_k ($k=1, 2, \dots, n$)，为网络处理后的期望值，激活函数一般选择线性传递函数， $W_{jk}^{(2)}$ ($j=1, \dots, l; k=1, \dots, n$) 为隐藏层和输出层的连接权值。

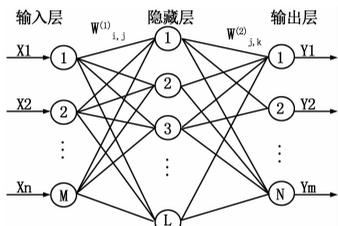


图 1 三层 BP 网络结构

1.2 BP 神经网络的反馈传播算法

对于 BP 网络，如何求得满足特定需求的权值集合 W 至关重要。假设对于输入样本 S 的理想输出为 $\{p_i^s\}$ ，实际输出为 $\{o_i^s\}$ ，则：

$$E_w^s = \frac{1}{2} \sum_i (p_i^s - o_i^s)^2 \tag{1}$$

度量了在一组给定权值下，实际输出与理想输出之间的差距，它与网络权值向量 W 的选取密切相关。因此问题可以归纳为如何调整网络权值得到 (1) 式的极小值。BP 网络的学习过程采用反向传播算法，通过不断调整网络权值 W 使得差值 E_w 达到理想值，从而使得 BP 网络预测模型达到理想精度。实际应用中，可以通过大量样本数据对网络进行训练，以达到较高精度要求。

1.3 BP 神经网络输入预处理

为了避免 BP 网络输入矢量中各维数据间因数量级差别而造成的误差，提高网络预测能力，在网络训练前需对样本数据进行必要的归一化处理。本文使用的归一化方法为：待输入数据向量减去历史数据均值，再除以标准差。对于确定的实验环境，均值和标准差可以通过大量的历史数据分析得出。假设在某时刻 t ，传感器的测量矢量为 (x'_1, \dots, x'_m) ，按上述方法经 (2) 式归一化处理后的到 BP 网络的输入矢量 $(\bar{x}'_1, \dots, \bar{x}'_m)$ 。

$$\bar{x}'_i = \frac{x'_i - \mu_i}{\sigma_i} (i = 1, \dots, m) \tag{2}$$

式中， μ_i 为 BP 网络第 i 维输入端历史数据均值； σ_i 为 BP 网络第 i 维输入端历史数据标准差。

2 动态采样驱动模型

2.1 节点参数间相关性分析

实验中无线传感器节点监测参数包括环境温度、空气相对湿度、土壤温度。这些参数在时间上往往具有连续性和周期性，图 2 显示了相同时间段内它们的典型变化。可以看出土壤温度和环境温度、空气相对湿度具有显著的相关性，且它们在时间序列上具有相似的周期性。

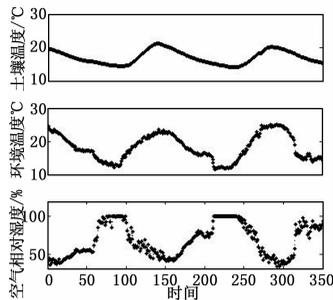


图 2 空气温度、湿度和土壤温度典型变化

选取 2013 年 5 月 10 日~2013 年 5 月 12 日北京市小汤山国家精准农业示范基地 300 组数据，每组包括环境温度、空气相对湿度和土壤温度 3 个变量。由表 1 可知土壤温度和环境温度、空气相对湿度之间的相关性较为突出，分别到达了 0.932 和一 0.830。因此将环境温度、空气相对湿度作为 BP 网络输入，土壤温度为输出是合理的。

表 1 参数相关性表(样本数 $n=300$)

相关性	空气相对湿度	环境温度
土壤温度	-0.830	0.932

2.2 阈值模型分析

周期性数据曲线可以分为平滑段和波动段，平滑段曲线可以用一元线性回归方程近似描述还原，该段节点用较低的采样率即可获得较高的数据还原度。对于波动段，节点必须通过更高的采样频率，获得足够数据点才能较好的展现参数的真实变化情况，如图 3 所示。对于数据曲线，主要关注的是其波动段，因为波动段表明外界环境变化最为剧烈，对于决策方案的设计最为重要。

分析环境历史数据 (图 2 所示) 可知，对于具体季节 (如夏季)，监测点数据具有平稳的极值点和周期性，数据波动段均处于极值点附件的区域内。因此可以通过分析历史数据中极值来划定波动段区间 (设为阈值区间)，通过判断 BP 网络预测输出值是否处于阈值区间来调整节点采集频率。其中阈值根据不同季节或时段变化从历史数据中分析得出。

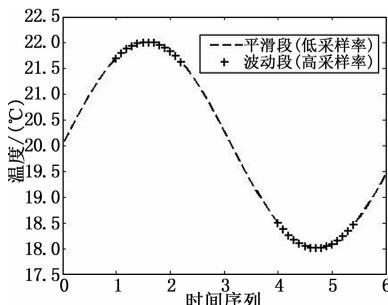


图 3 数据曲线平滑段与波动段

假设从历史数据中统计得出某时间段内数据极大值、极小值向量分别为 $(max_1, max_2, \dots, max_m)$, $(min_1, min_2, \dots, min_m)$, 计算其均值 \overline{max} , \overline{min} 。由此可以得到阈值模型区间为:

$$\begin{aligned} Threshold_{max} &= \overline{max} \pm \epsilon_1 \\ Threshold_{min} &= \overline{min} \pm \epsilon_2 \end{aligned} \quad (3)$$

式中, ϵ_1 , ϵ_2 为服从正态分布的浮动区间, 它们的值可以根据需求通过实验得出。

2.3 动态采样算法

系统开始时刻, 节点的环境温度、空气相对湿度采样值经由 BP 网络处理后得到土壤温度的预测输出值, 若该值进入阈值模型 $Threshold_{max}$ 或 $Threshold_{min}$ 区间范围则判定土壤温度进入波动段, 节点增加土壤温度采样频率 (设为 $rate_{max}$), 否则视为平滑段, 降低采样率 (设为 $rate_{nor}$)。动态采样算法过程描述如下。

step 1: 选取样本数据训练 BP 网络。

step 2: 根据历史数据确定波动段范围 $Threshold_{max}$ 和 $Threshold_{min}$ 。

step 3: 节点以 $rate_{max}$ 进行数据采样, 根据环境温度、空气相对湿度采样值经 BP 网络处理后得到土壤温度预测值 $soilTemp$ 。

step 4: 如果 $soilTemp$ 进入区间 $Threshold_{max}$ 或 $Threshold_{min}$ 转 step 3, 否则转 step 5。

step 5: 节点以 $rate_{nor}$ 进行数据采样, 根据环境温度、空气相对湿度采样值通过 BP 网络预测得到土壤温度值 $soilTemp$, 转 step 4。

为保证算法准确性, 根据不同季节变化, 阈值区间需根据历史数据作相应的调整。

3 实验分析

3.1 BP 网络训练结果与预测结果分析

采用 Matlab 平台进行数据仿真, 5 月 10~11 日的 200 组作为训练数据, 5 月 12 日的 100 组作为预测数据。该段数据变化稳定, 具有很高的代表性。

BP 网络隐含层激活函数采用 sigmoid 非线性 tansig 函数, 输出层激活函数采用线性的 purelin, 训练函数采用有动量的梯度下降法 traingdm。训练迭代次数设定为 5 000 次, 理想误差设置为 0.01。为了验证模型精确度, 对模型进行误差分析, 分别采用回归估计均方根误差 RMSE (root-mean-square error) 和绝对误差 AE (absolute error) 对预测值和实测值进行统计分析。计算公式分别为:

$$\begin{aligned} RMSE &= \sqrt{\frac{\sum_{i=1}^n (act_i - fct_i)^2}{n}} \\ AE &= \frac{1}{n} \sum_{i=1}^n |act_i - fct_i| \end{aligned} \quad (4)$$

式中, act_i 为实际观测值; fct_i 为预测值; n 为采样次数

通过多次试验, 对比不同隐藏层节点数及相应的均方根误差 (见表 2), 最终确定 BP 网络采用 2-9-1 的结构。

对网络训练结果进行反归一化将数据映射至正常范围内, 然后作土壤温度训练值与实际值拟合曲线, 如图 4 所示。可以看出, 训练值与实测值基本集中在 1:1 线附近, 决定系数为 0.930, 网络训练结果是令人满意的。其中 RMSE 和 AE 分别达到 0.53 °C, 0.40 °C, 网络的预测精度较高。

表 2 神经网络隐藏层节点数与相应训练误差

隐层节点数	RMSE
8	0.1006
9	0.0702
10	0.0741
11	0.0731
12	0.0705

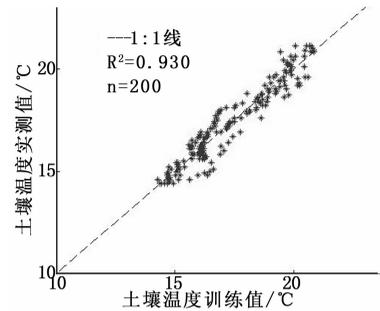


图 4 土壤温度训练值、实测值对比图

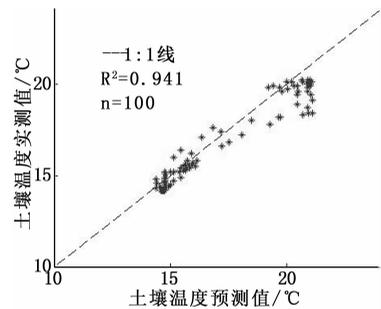


图 5 土壤温度预测值、实测值对比

网络的预测结果如图 5 所示, 预测数据为 100 组, 结果显示土壤温度实测值与预测值决定系数为 0.941, BP 网络学习效果较好。但其中也存在一些噪声较大的点, 最大误差为 3.90 °C, 可能是由于监测环境受到人为因素的干扰 (如田间灌溉)。预测结果误差 RMSE 和 AE 分别达到 0.83 °C, 0.54 °C, 预测值和实测值趋势基本一致。

3.2 节点调度算法性能分析

依据动态采样算法描述, 对预测数据进行处理。分析历史数据, $Threshold_{max}$ 取值 20.07 °C, $Threshold_{min}$ 取值 14.03 °C。根据实际试验精度要求, ϵ_1 、 ϵ_2 分别取值 1.1 °C 和 0.6 °C。采样频率 $rate_{max}$ 为 10 分钟一次, $rate_{nor}$ 为 30 分钟一次。为了便于观看, 在动态采样结果曲线上增加幅值 (本文取 2 °C), 结果如图 6 所示。

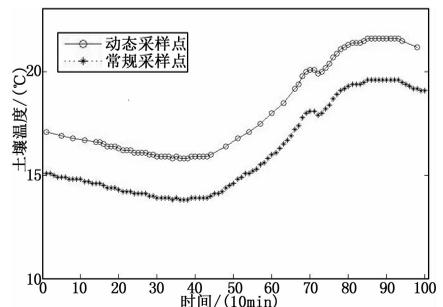


图 6 动态-常规采样对比图

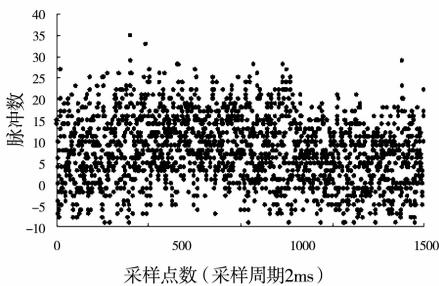


图 7 电机 1 与电机 3 的实时位置误差

通过以上实验验证得知, 采用包米勒 Controller PLC02 基于其 EtherCAT 通信模块设计的具有多级主从式结构的造波机同步运动控制系统在网络实时性、电机同步性、运行稳定性等方面完全满足造波机的应用, 且提高了造波机的同步控制精度。

5 结语

研究了基于包米勒 Controller PLC02 和其通信模块组成的基于 EtherCAT 网络的具有多级主从式结构的同步运动控制系统, 并对其在分布式 156 轴蛇形造波机应用中存在的问题提出了具体解决方案。通过大量的试验验证了这种多控制器同步运动控制系统设计在蛇形造波机应用中的可行性, 实现了系统各轴严格意义上的精确同步控制。为研制更大型的造波机系统提供了技术支持, 并且也为其它领域的大型复杂运动控制系统的

(上接第 2487 页)

从图 6 中可以看出, 动态采样算法可以很好的区分数据平滑段和波动段, 平滑段采用较小的采样频率, 当进入波动段时增加采样频率, 因此算法还原出的数据曲线与实际曲线基本一致。在常规采样 100 次的条件下, 动态采样降至 71 次, 数据减少近 30%, 很好的减少了系统的数据冗余和传输量。

4 结论

针对农业环境监测传感器网络应用中, 参数变化具有较强的周期性和内在耦合、关联性的特点, 构建了一种基于误差反向传播的多层前馈神经网络预测和阈值分析的传感器网络动态采样模型, 并以环境温度、空气相对湿度、土壤温度 3 个典型参数为实验对象, 构建了其基于 BP 神经网络的参数动态采样调度模型。

仿真实验结果表明, BP 神经网络模型预测值和实测值之间的均方根误差 RMSE 及绝对误差 AE 分别为 0.83 °C 和 0.54 °C。按照仿真实验得出的最佳阈值模型设定值, 相比于连续采样, 动态采样次数减少 30%, 模型应用效果较显著。由于受时间限制, 调度模型还仅局限于研究环境温度、空气相对湿度等对土壤温度的影响, 还未能综合考虑光照、光合有效辐射、土壤湿度等参数的相互影响。下一步, 将通过研究更多参数历史数据, 分析其内在相关性, 建立合适的预测方法, 从而建立更多参数的动态采样驱动模型。

参考文献:

[1] 陈丹, 郑增威, 李际军. 无线传感器网络研究综述 [J]. 计算机测量与控制, 2004, 12 (8): 701-704.
 [2] 李凤保, 李凌. 无线传感器网络技术综述 [J]. 仪器仪表学报, 2005, 26 (8): 559-561.
 [3] 蒲荣富. 大规模无线传感器网络能效性分析 [J]. 电子测量与仪器

同步设计提供了参考。

参考文献:

[1] Romijnders L N G. The Development of a New Segmented Deepwater Wave Generator [A]. Ocean Wave Measurement and Analysis (2001) [C]. ASCE, 2002; 1209-1217.
 [2] 李木国, 徐福荣. 基于 SynqNet 的多向不规则波造波机控制系统 [J]. 测控技术, 2007, 26 (10): 37-39.
 [3] 张鹏, 魏仁哲. 基于 TRIO 运动控制器的 144 轴运动控制系统 [J]. 控制工程, 2009, 16 (s): 21-23.
 [4] 李木国, 陈永富, 王静. 多控制器多轴同步技术用于造波机系统的研究 [J]. 计算机测量与控制, 2013, 21 (11): 2985-2987.
 [5] Marin R, Joseph E, Stubbs, Dmitry D. EtherCAT Enabled Advanced Control Architecture [C]. San Francisco, CA: IEEE, 2010; 9-44.
 [6] Baumuller Nuremberg GmbH. Application Manual for b maXX controller PLC [Z]. 2007.
 [7] Prytz G., Skaalvik J. Redundant and synchronized EtherCAT network [C]. Trento, Italy: IEEE, 2010; 201-204.
 [8] 汪雅楠, 谭南林. EtherCAT 时钟同步的误差仿真分析 [J]. 自动化技术与应用, 2011, 30 (7): 22-26.
 [9] GmbH B. A. Hardware Data Sheet ET1100 - EtherCAT Slave Controller [Z]. 2008.
 [10] 李木国, 柳淑学, 张群, 等. 蛇形多向不规则波造波机波浪产生方法及特性 [J]. 大连理工大学学报, 2003, 43 (3): 354-361.
 [11] 李木国, 柳淑学, 张群, 等. 蛇形多向不规则波造波机波浪产生学报, 2007, 21 (6): 61-65.
 [12] 孙利民, 李建中, 陈渝. 无线传感器网络 [M]. 北京: 清华大学出版社, 2005.
 [13] Goldsmith A J, Wicker S B. Design challenges for energy-constrained Ad Hoc wireless networks [J]. IEEE Wireless Communications, 2002, 9 (4): 8-27.
 [14] 李建中, 石胜飞, 王朝坤. 基于感知数据概率模型的无线传感器网络采样和通信调度算法 [J]. 计算机应用, 2005, 25 (9): 1982-1985.
 [15] 李珂, 贾惠波, 吴晓培, 等. 用于土壤湿度测量的传感器网络动态采样策略 [J]. 电子测量与仪器学报, 2011, 25 (4): 338-342.
 [16] Shuman, A. Nayyar, A. Mahajan, Y. Goykhman, et al. Measurement Scheduling for Soil Moisture Sensing: From Physical Models to Optimal Control, in Proceedings of the IEEE Special Issue on Sensor Networks and Applications [J]. 2010, 98 (11): 1918-1933.
 [17] 牛建军, 邓志东. 基于马尔可夫链的无线传感器网络分布式调度方法 [J]. 自动化学报, 2010, 36 (5): 685-695.
 [18] Hornik K M, Stinchcombe M, White H. Multilayer feed forward networks are universal approximators [J]. Neural Networks, 1989, 2 (2): 359-366.
 [19] 张绍秋, 胡跃明. 基于 BP 神经网络的税收预测模型 [J]. 华南理工大学学报: 自然科学版, 2006, 34 (6): 55-58.
 [20] 张学工. 关于统计学习理论与支持向量机 [J]. 自动化学报, 2000, 26 (1): 32-43.
 [21] Conejo A J, Plazas M A, Espinola R, et al. Day-ahead electricity price forecasting using the wavelet transform and ARIMA model [J]. IEEE Trans Power System, 2005, 20 (2): 1035-1042.
 [22] 王军, 吕震中, 曹荃. 基于神经网络的传感器故障诊断方法与应用 [J]. 自动化仪表, 2003, 24 (10): 21-24.
 [23] 毛颖科, 关志成, 王黎明. 基于 BP 人工神经网络的绝缘子泄漏电流预测 [J]. 中国电机工程学报, 2007, 27 (27): 7-12.